

一种基于扩展颜色 Petri 网的 Web 服务组合验证机制

李景霞 闫春钢

(同济大学计算机科学与工程系 上海 201804)

(同济大学嵌入式系统与服务计算教育部重点实验室 上海 201804)

摘要 Web 服务技术日趋成熟,它提供了一种平台独立、自描述、位置透明的软件模块。利用 Web 服务组合技术,能够快速、灵活地满足复杂多变的业务需求。随着 Web 服务组合流程日趋复杂,组合流程设计越来越容易出错。给出了一个描述 Web 服务组合流程的扩展层次颜色 Petri 网模型,它独立于具体的流程描述语言,支持流程的层次表达且能较全面地刻画 Web 服务组合流程。还给出了模型的控制流网正确性、数据流网正确性、流程实例正确性的验证方法,从 3 个层面确保了 Web 服务组合流程的正确性。

关键词 层次颜色 Petri 网,模型驱动架构,Web 服务组合,正确性分析

中图分类号 TP393 **文献标识码** A

Verification Mechanism for Web Service Composition Based on Extended Colored Petri Net

LI Jing-xia YAN Chun-gang

(Department of Computer Science & Engineering, Tongji University, Shanghai 201804, China)

(Key Laboratory of Embedded System and Service Computing of Ministry of Education, Tongji University, Shanghai 201804, China)

Abstract The development of Web service technology provides us a kind of platform-independent, self-described, location-transparent software module. Utilizing Web service composition technology, business demand can be met quickly and flexibly. As Web service composition becomes more and more complexity, design of composition process became more and more error-prone. We put forward a model for Web service composition description based on extended colored Petri net. This model is independent of any process description languages, supports hierarchical process description and can describe a more comprehensive process. Moreover, we also presented methods to validate control flow correctness, data flow correctness and process instance correctness of the process, thus we can insure the process correctness from these three levels.

Keywords Hierarchical colored Petri net, Model driven architecture, Web service composition, Correctness analysis

1 引言

随着 Internet 在全球的发展,基于网络的分布式应用越来越呈现出多样性和复杂性。Web 服务提供了一种分布的、自描述的、模块化的组件,遵循一定的技术规范,执行特定的任务,为基于组件的系统开发提供了一种平台独立的组件。但单个 Web 服务能提供的功能有限。有必要将共享的 Web 服务组合起来,提供功能更为强大的服务,促使 Web 服务组合研究的兴起。

当前出现了诸多的过程定义描述语言,如 BPEL4WS^[1], BPML^[2], WSCI^[3], OWL-S^[4]等,从不同角度出发给出了 Web 服务组合的语法规则。但这些语言仅给出了基于 XML 过程定义的语法表达,不能分析所描述组合服务流程的正确性。实践证明,定义和模拟真实世界中的业务流程,是一个非常复杂且容易出错的过程。有错的流程定义将导致错误的实现,

在系统运行之后再更正错误,耗费巨大。现阶段不少研究工作^[5-8]是针对这一问题展开的,对具体的流程描述语言给出与某个形式化模型的映射法则,将流程描述转换成形式化描述,进而进行验证。然而,这些验证工具均针对具体的流程描述语言,依赖于具体的流程描述语言。随着描述语言的改进,需要修改映射规则。而对新提出的描述语言,则需要开发新的映射规则,因此应用范围比较窄。本文借鉴 MDA 思想,给出了一种用于描述 Web 服务组合的扩展层次颜色 Petri 网模型,该模型独立于具体的流程描述语言,支持流程的层次描述且能较全面地刻画流程,同时支持流程的正确性分析。

2 WSC_EHCPN 模型定义

一个描述 Web 服务组合的扩展层次颜色 Petri 网^[9]模型 WSC_EHCPN 是一个三元组, $WSC_EHCPN = (EHCPN, T_i, PA)$, 其中

到稿日期:2008-11-19 返修日期:2009-02-16 本文受国家“八六三”高技术研究发展计划项目“可靠的服务组合方法及其验证技术研究”(2007AA01Z136)和国家“九七三”重点基础研究发展规划项目“语义网格的过程语义及其验证理论与服务集成技术”(2003CB316902)资助。

李景霞(1976-),女,博士,主要研究方向为 Web 服务组合、服务计算, E-mail: jxiali@163.com; 闫春钢 博士,教授,主要研究方向为 Petri 网理论、工作流建模方法、Web 服务。

(1) $EHCPN(\Sigma, P_c, P_d, T, A_c, A_d, C, G, E_c, E_d, AT, I)$ 是一个非层次的扩展 HCPN, 其中

(1.1) Σ 是 Web 服务组合中涉及的数据类型集合, 包括数据托肯和控制托肯, 控制托肯类型 $CONTROL \in \Sigma$;

(1.2) P_c 为控制托肯的非空有限库所集, 是存放 Web 服务控制参数的缓冲区;

(1.3) P_d 为数据托肯的非空有限库所集, 是存放 Web 服务数据参数的缓冲区;

(1.4) $T = T_c \cup T_0 \cup T_s$ 为非空有限变迁集。其中 T_c 是普通变迁集, 是调用的 Web 服务集合; T_0 是零变迁集, $\forall t \in T_0$ 不关联到任何 Web 服务, 不执行任何处理, 仅辅助流程的表达; T_s 是替代变迁集, 是调用的 Web 服务组合子流程集和 Web 服务循环执行流程集的并集, 每个替代变迁对应一个 WSC_EHCPN;

(1.5) $A_c \subseteq P_c \times T \cup T \times P_c$ 是一有限控制弧集, 且 $P_c \cap T = P_c \cap A_c = T \cap A_c = \phi$;

(1.6) $A_d \subseteq P_d \times T \cup T \times P_d$ 是一有限数据弧集, 且 $P_d \cap T = P_d \cap A_d = T \cap A_d = \phi$;

(1.7) C 是一数据类型函数, 定义为 $C: (P_c \cup P_d) \rightarrow \Sigma$, 即指定库所 $p \in (P_c \cup P_d)$ 中的参数类型为 $C(p)$;

(1.8) G 是一防卫函数, 定义为 $G: T \rightarrow G(t)$, 满足 $\forall t \in T: [Type(G(t)) = B \wedge Type(Var(G(t))) \subseteq \Sigma]$, 用于指定调用 Web 服务必须满足的前提条件;

(1.9) E_c 是控制弧 A_c 上的函数, $\forall a_c \in A_c$ 有 $E_c(a_c) = c$, E_c 表示 Web 服务控制参数的输入、输出, 表示 Web 服务调用完成与否;

(1.10) E_d 是数据弧 A_d 上的函数, 定义为 $E_d: A_d \rightarrow E_d(a_d)$ 。满足 $\forall a_d \in A_d: [Type(E_d(a_d)) = C(p_d)_{MS} \wedge Type(Var(E_d(a_d))) \subseteq \Sigma \setminus CONTROL]$, 其中 p_d 是 $A_d(a_d)$ 中的库所, $C(p_d)_{MS}$ 是库所 p_d 的数据类型集上的多重集, E_d 用于表示 Web 服务数据参数的输入、输出;

(1.11) AT 是变迁属性函数, 定义为 $AT: T \rightarrow Attributes(tType, tName, inputSet, outputSet)$, 为每一 Web 服务指定变迁类型、名称、输入/出参数集。变迁类型有 3 种: 普通(common)、零变迁(zero)、替代变迁(substitute);

(1.12) I 是一初始化函数, 定义为从 $(P_c \cup P_d)$ 到一个封闭表达式 $I(p)$, 满足 $\forall p \in (P_c \cup P_d): [Type(I(p)) = C(p)_{MS}]$;

(2) T_s 是替代变迁集, $\forall t \in T_s: t = (WSC_EHCPN, P_p, tsType)$, 含 t 的网称为超页, t 对应的网称为子页, 其中

(2.1) T_s 是调用的 Web 服务组合子流程集和 Web 服务循环执行流程集的并集, 用于复用已定义好的 Web 服务组合或实现 Web 服务的循环执行; 由 T_s 定义可见, WSC_EHCPN 是嵌套定义的, 若 $T_s = \phi$, 则 WSC_EHCPN 退化为非层次 WSC_EHCPN;

(2.2) P_p 是端口库所集, 是子页与超页的耦合库所;

(2.3) $tsType$ 是替代变迁类型函数, 定义为 $tsType \rightarrow \{subProcess, Iterate\}$ 。有两类替代变迁: 表示子流程的和表示循环结构的。对于子流程的表达, 以替代变迁及相应的子网来表示, 子流程在超页中表现为一个替代变迁; 对于循环结构的表达, 以替代变迁及相对应的子网来表示, 被循环调用的 Web 服务在超页中表现为一个替代变迁;

(3) PA 是端口分配函数, 用于关联超页的套接库所和子页的端口库所, 定义为从 T_s 到二元关系:

(3.1) $\forall t \in T_s: PA(t) \subseteq (\cdot t \cup t') \times t, P_p, t, P_p$ 表示替代变迁 t 的端口库所集;

(3.2) $\forall t \in T_s, \forall (p_1, p_2) \in PA(t): [C(p_1) = C(p_2) \wedge I(p_1) \langle \rangle = I(p_2) \langle \rangle]$ 。

对 WSC_EHCPN 网模型, 定义流程入口 $P_i = \{p_i | p_i \in P_c \wedge \cdot p_i = \phi\}$ 、流程出口 $P_o = \{p_o | p_o \in P_c \wedge p_o' = \phi\}$ 。对所有 $|P_i| > 1$ 的流程, 都能通过添加一个控制库所、一个零变迁, 一条控制库所到零变迁的有向弧并标识该弧的函数为一个控制托肯, 从零变迁分别到原流程每个入口库所的有向弧并标识这些弧的函数为一个控制托肯, 即能在不改变流程逻辑结构的情况下, 将流程改造成 $|P_i| = 1$, 设定 $P_i = \{p_i\}$ 。

3 Web 服务基本组合结构的 WSC_EHCPN 表示

3.1 原子 Web 服务的 WSC_EHCPN 表示

原子 Web 服务是一个基本 Web 服务, 如图 1 所示, 没有调用其他 Web 服务, 表示为一个非层次的扩展颜色 Petri 网, Web 服务是网中的变迁, 控制参数和数据参数分别存放在控制库所、数据库所中。

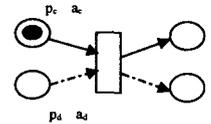


图 1 原子 Web 服务

3.2 基本 Web 服务组合的 WSC_EHCPN 表示

基本的 Web 服务组合结构包括顺序、并行、选择、循环, 其 WSC_EHCPN 表示如图 2—图 5 所示。图中灰色变迁是零变迁, 不映射到任何 Web 服务, 仅辅助流程表达。

1) 顺序: 两个 Web 服务顺序执行。

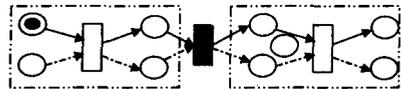


图 2 Web 服务顺序组合

2) 并行: 两个 Web 服务并行执行。

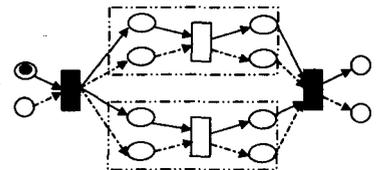


图 3 Web 服务并行组合

3) 选择: 两个 Web 服务根据条件选择执行。

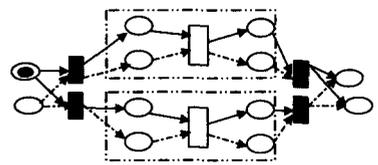


图 4 Web 服务选择组合

4) 循环: 一个 Web 服务循环执行。

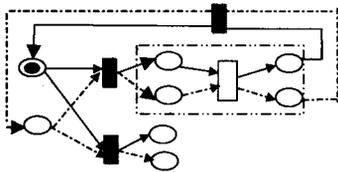


图5 Web 服务循环组合

4 Web 服务组合流程的分析验证

利用形式化模型来描述 Web 服务组合模型,其目的是辅助流程的设计、分析和验证。下面给出基于该模型的 Web 服务组合的分析验证。

4.1 控制流网的正确性

对 WSC_EHCPN 模型,抽取其中的控制流关系,得到 Web 服务组合流程的控制流网 $WSC_HC_{on}PN(P_c, T, A_c, E_c)$,描述组合流程包含的 Web 服务及 Web 服务间的依赖关系,是所有流程实例共享的流程结构。在对 WSC_ECPN 模型的验证中,首先要确保 $WSC_HC_{on}PN(P_c, T, A_c, E_c)$ 是正确的,以保证流程的结构正确性。借鉴工作流模型合理性定义^[10],一个 WSC_HConPN 网是正确的当且仅当:

1)对每个初始标识 M_0 可达的状态 M ,必存在变迁引发序列,使得网标识从 M 转换为 M_c (只有某个出口库所有控制托肯,其他任何控制库所都没有控制托肯),即 $\forall M(p_c \xrightarrow{*} M) \Rightarrow (M \xrightarrow{*} p_c)$;

2)标识 M_c 是一组正常结束标识,即 $\forall M(p_c \xrightarrow{*} M \wedge M \geq p_c) \Rightarrow (M = p_c)$;

3)不存在死变迁,即 $\forall t \in T, \exists M, M_1$,使得 $(p_c \xrightarrow{*} M \xrightarrow{t} M_1)$ 。

对控制流网的正确性验证可以使用 Petri 网的分析方法^[11]加以分析验证。

4.2 数据流网的正确性

WSC_EHCPN 模型的数据流反映出服务组合中 Web 服务操作间的数据依赖关系。对结构正确的 Web 服务组合流程,参照控制流网来检测数据流网与控制流网是否存在冲突。

定义 1(控制依赖) 在 WSC_EHCPN 网中,如果从节点 i 到节点 j 有一条路径,且组成该路径的所有弧都为控制弧,则称 j 控制依赖于 i ,记为 $jCDi$ 。

定义 2(数据依赖) WSC_EHCPN 网中,若从节点 i 到节点 j 有一条路径,且组成该路径的每条弧都是数据弧,则称 j 数据依赖于 i ,记为 $jDDi$ 。

定义 3(依赖) 在 WSC_EHCPN 中,如果有一条路径 $Path = \langle i = v_1, \dots, v_i, \dots, v_n = j \rangle$,且每个节点 v_{i+1} 都控制依赖或数据依赖于 v_i ,称 j 依赖于 i ,记为 jDi 。

定义 4(控制依赖关系是传递的) 在 WSC_HConPN 中,在集合 $P_c \cup T$ 上,有 $jCDi$ 成立,表示节点 j 只有在节点 i 完成才有可能执行。对任意节点 i, j 和 k 有控制依赖关系 $jCDi, kCDj$,可以得出节点 k 是控制依赖于节点 i 的,即有 $kCDi$ 。

同理可以得出数据依赖关系是传递的。根据依赖闭包的求解算法^[12]分别求解出控制依赖闭包、数据依赖闭包和依赖闭包,检测数据依赖闭包及依赖闭包与控制依赖闭包是否存

在冲突。如果没有冲突,则数据流网是正确的。

4.3 流程实例的正确性

在 WSC_EHCPN 模型中,控制流网给出了 Web 服务组合流程的结构图,具体的流程执行案例则由控制流和数据流共同决定。

定义 5(变迁绑定) 是定义在 $Var(t)$ 之上的函数 b :

- 1) $\forall v \in Var(t); b(v) \in Type(v)$;
- 2) $G(t) \langle b \rangle$ 为“真”。

变迁 t 的绑定是把变迁 t 的每个变量 $Var(t)$ 用合适的数值代替,且使得防卫表达式 $G(t) \langle b \rangle$ 为真,即绑定一定满足防卫表达式。

定义 6(绑定元素) 是一个 (t, b) 对,其中 $t \in T, b \in B(t)$ 。

定义 7(步的使能) 当 WSC_EHCPN 模型处于一个标识 M 时,称步 Y 是使能的,当且仅当满足:

$$(\forall p_d \in P_d: \sum_{(a,b) \in Y} E_d(p_d, t) \langle b \rangle \leq M(p_d)) \wedge (\forall p_c \in P_c \wedge p_c \in \cdot t: \sum E_c(p_c, t) \leq M(p_c))$$

此时称绑定 (t, b) 有效,也称变迁 t 可引发。

定义 8(步发生后的标识) 当在标识 M_1 时步 Y 使能,则其发生后标识变为 M_2 :

$$\forall p_d \in P_d, M_2(p_d) = (M_1(p_d) - \sum_{(a,b) \in Y} E_d(p_d, t) \langle b \rangle) + \sum_{(a,b) \in Y} E_d(p_d, t) \langle b \rangle$$

$$\forall p_c \in P_c, M_2(p_c) =$$

$$\begin{cases} M_1(p_c) - E_c(p_c, t) & p_c \in \cdot t - t \cdot \\ M_1(p_c) + E_c(t, p_c) & p_c \in t \cdot - \cdot t \\ M_1(p_c) - E_c(p_c, t) + E_c(t, p_c) & p_c \in \cdot t \cap t \cdot \\ M_1(p_c) & p_c \notin \cdot t \cup t \cdot \end{cases}$$

通过库所的实例化函数及变迁绑定将流程实例化,通过上述的变迁引发规则,可以对流程实例的运行进行仿真,检测流程实例的正确性。

结束语 针对当前 Web 服务组合建模不独立于具体的流程语言且不能全面刻画 Web 服务组合流程的问题,本文给出了一种扩展层次颜色 Petri 网模型 WSC_EHCPN,用于描述 Web 服务组合,并给出了 4 种基本 Web 服务组合结构的 WSC_EHCPN 表示。接着从控制流网正确性、数据流网正确性、流程实例正确性 3 个层面验证了 Web 组合流程的正确性。给出形式化的 Web 服务组合描述模型,只是研究工作的一个开端。今后的研究工作将在如下几个方面展开:WSC_EHCPN 模型到 Web 服务组合描述语言的自动转换、Web 服务组合流程的优化等。

参考文献

- [1] Andrews T, Curbera F, Dholakia H, et al. Business Process Execution Language for Web Services Version 1. 1. Specification [S]. 2003
- [2] Arkin A. Business Process Modeling Language (BPML) [S]. Spec., BPML org, 2002
- [3] Arkin A, Askary S, Fordin S, et al. Web Service Choreography Interface (WSCI) 1. 0 [OL]. <http://www.w3.org/TR/wsci/>
- [4] Martin D, Ankolekar A, Burstein M, et al. OWL-S 1. 1 Release [OL]. <http://www.damll.org/services/owl-s/1.1>
- [5] Koshkina M, van Breugel F. Modelling and verifying web service

orchestration by means of the concurrency workbench[J]. ACM SIGSOFT SEN, 2004, 29(5): 1-10

[6] Hamadi R, Benatallah B. A Petri Net - based Model for Web Service Composition[C]//Proceedings of the Fourteenth Australasian Database Conference on Database Technologies 2003. Adelaide, Australia, February 2003; 191-200

[7] Tao Xiaofeng, Jiang Changjun. Formalizing Web Service and Modeling Web Service-based System Based on Object Oriented Petri Net[J]. GCC, 2003(1): 1008-1011

[8] A CPNets-based Design and Analysis Framework for Service Oriented Distributed Systems[D]. USA: University of Georgia, 2005

[9] Jensen K. Colored Petri Nets; Basic Concepts, Analysis Methods

and Practical Use[M]. Volume 1; Basic Concepts. Monographs in Theoretical Computer Science. 2nd corrected printing, Berlin, Heidelberg, New York; Springer-Verlag, 1997

[10] van der Aalst W M P. Making Work Flow; On the Application of Petri nets to Business Process Management[C]//Esparza J, Larkos C, eds. Application and Theory of Petri Nets 2002. volume 2360 of Lecture Notes in Computer Science. Berlin; Springer-Verlag, 2002; 1-22

[11] Murata T. Petri nets; Properties, analysis and applications[J]. Proc. IEEE, 1989, 77(4): 541-580

[12] 左孝凌, 李为鑑, 刘永才. 离散数学[M]. 上海: 上海科学技术文献出版社, 1982

(上接第 123 页)

6.2.2 延迟-代价积

图 5 示出了每个算法在分布式构建广播树时所产生的延迟-代价积。

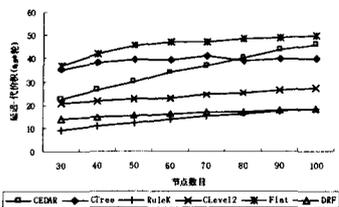


图 5 延迟-代价积

如图 5 所示, 参加比较的 6 种算法按照其延迟-代价积由小到大排序, 依次为 RuleK, DRF, CLevel2, CEDAR 和 Flat; CTree 处于 CLevel2 和 Flat 之间, 在 30 个节点到 70 个节点之间时高于 CEDAR, 80 个节点到 100 个节点之间时低于 CEDAR。在图 5 中, 虽然 DRF 算法的延迟-代价积平均高出 RuleK 算法 18.5%, 但是随着节点数目逐步增大, 两者在延迟-代价积方面的差距明显呈现出逐渐减小的趋势。另一方面, 图 4 显示, 在平均意义上, DRF 算法的广播延迟已缩短至 RuleK 算法的 1/2.8; 而且随着节点数目的逐步增大, 这一比例呈现继续减小的趋势。

结束语 尽管在单速率无线自组织网络中, 构建最小连通支配集是降低广播延迟的有效方法, 但是这一方法并不能直接用在多速率无线网状网中。本文通过引入三维邻接矩阵, 形式化地描述了多速率无线网状网中广播延迟最小化问题, 并提出了一种用于构建广播树的分布式速率优先算法。本算法的特点是每个节点以某个概率充当广播节点, 该概率与可供选择的传输速率和局部拓扑信息有关。大量仿真结果表明, 我们的算法与现有算法相比, 其广播延迟获得了显著降低。在实际工程应用中, 在多速率无线网状网中部属广播服务时, 给予较高速率以较高的优先选择权将是缩短全网广播延迟的有效途径。

参考文献

[1] Akyildiz I F, Melodia T, Chowdhury K R. A survey on wireless multimedia sensor networks[J]. Computer Networks, 2007, 51

(4): 921-960

[2] Hiertz G R, Max S, Zhao Rui, et al. Principles of IEEE 802. 11s [C]//Proc. of 16th IEEE Int'l. Conf. on Computer Communications and Networks (ICCCN). New York: IEEE Press, 2007; 1002-1007

[3] Katia O, Kumar V, Gene T. Flooding for reliable multicast in multi-hop Ad hoc networks[J]. Wireless Networks, 2001, 7(6): 627-634

[4] Tseng Yu-Chee, Ni Sze-Yao, Chen Yuh-Shyan, et al. The broadcast storm problem in a mobile Ad hoc network[J]. Wireless Networks, 2002, 8(3): 153-167

[5] Dai F, Wu J. Distributed dominant pruning in Ad hoc wireless networks[C]//Proc. of IEEE Int'l. Conf. on Communications (ICC). New York: IEEE Press, 2003; 353-357

[6] Wan P-J, Alzoubi K M, Frieder O. Distributed construction of connected dominating set in wireless Ad hoc networks[J]. Mobile Networks and Applications, 2004, 9(2): 141-149

[7] Sinha P, Sivakumar R, Bharghavan V. Enhancing ad hoc routing with dynamic virtual infrastructures[C]//Proc. of IEEE Joint Conf. of Computer and Communications Societies (INFOCOM). New York: IEEE Press, 2001; 1763-1772

[8] Dai F, Wu J. Virtual backbone construction in MANETs using adjustable transmission ranges[J]. IEEE Transactions on Mobile Computing, 2006, 5(9): 1188-1200

[9] Chou C T, Misra A, Qadir J. Low-latency broadcast in multirate wireless mesh networks[J]. IEEE Journal on Selected Areas in Communications, 2006, 24(11): 2081-2091

[10] Kuhn F, Wattenhofer R. Constant-time distributed dominating set approximation[J]. Distributed Computing, 2005, 17(4): 303-310

[11] Bauer J, Haugland D, Di Yuan. Analysis and computational study of several integer programming formulations for minimum-energy multicasting in wireless Ad hoc networks[J]. Networks, 2008, 52(2): 57-68

[12] 钱颂迪, 等. 运筹学(第 2 版)[M]. 北京: 清华大学出版社, 1990: 117-120

[13] Dai Fei. Wireless routing simulation suit (WRSS)[EB/OL]. <http://venus.ece.ndsu.nodak.edu/~fdai/>